**Quantitative HSQC Analyses of Lignin: A Practical Comparison**

Marco Sette *, Heiko Lange *, Claudia Crestini a*

**Abstract:** Lignin is the second-most abundant polymer after cellulose within the biomass of our planet. Structurally, it displays random oligomeric units without fixed repetition schemes beyond the stage of dimers. Quantitative \(^1\)H-\(^1\)C HSQC measurements have recently greatly facilitated lignin analyses. In some cases, however, long acquisition times needed for obtaining quantitative HSQCs are not compatible with the chemical integrity of (a potentially functionalised) lignin sample. We thus compared different methods that were developed for more time-efficient quantitative HSQC measurements with respect to their usefulness in lignin analyses: reliable and reproducible results were obtained using both the QQ-HSQC and the HSQC method.

3rd Annual User Group Meeting of BioNMR 2013

**Introduction**

‘Wasting’ or ‘leaving unused’ abundant renewable resources is a problem in light of the no longer justifiable use of fossil-based resources with respect to practical, ecological, and socio-economic reasons. Lignin from forest-biomass is one of the naturally abundant resources that are currently seen as waste. Serving both as stabilising and protecting element within the plants, where it is chemically ‘fixed’ in lignin-carbohydrate complexes (LCC), its non-uniform distribution within the plant cell walls sums up to species-specific amounts \(^1\) ranging from 20 ±4% in hardwoods, to 28 ±3% in softwoods and herbaceous angiosperms; monocots generally contain less lignin (15±4%). Obtained mainly as a by-product in the pulp- and paper-industries, its intriguingly diverse chemical structure does represent so far a major obstacle with respect to its widespread valorisation beyond the uses as energy source or as substrate for simple bulk chemicals such as vanillin or dimethyl sulfoxide (DMSO) \(^1\). For solving this dilemma, reliable, easy-to-use, and quick methods are needed that will allow the elucidation of the structural features of a given lignin, since only a sound structural information will guarantee, i.e., will allow the design of a (chemical) valorisation strategy that can account in detail for the lignin-typical complexity and heterogeneity.

One of the most reliable methods for investigating structural features of substances, and thus also of lignin, is NMR spectroscopy. Thanks to the advances in the technology as such, it is nowadays possible to gain structural information on lignin at an unprecedented accuracy level, both qualitatively and quantitatively; several articles and scientific monographs on this topic appeared in archival literature in the last years \(^4,5,6\). As it has been pointed out before, among others, two-dimensional heteronuclear NMR spectroscopy advanced the understanding of structural features of lignin, since especially qualitative and quantitative HSQC-analyses shed light on abundances of the different typical interunit bonding units in lignin.

Lignin (1) is a complex polyphenolic oligomer / polymer, displaying always a plant-specific composition \(^7,8\). To the best of the current knowledge, lignin lacks a – in terms of classical polymer-chemistry – defined primary structure, but rather represent random phenyl-propanoid (C9) polyphenols, which are mainly linked by arylglycerol ether bonds between phenolic para-coumaryl alcohol (2), (H-type), coniferyl alcohol (3) (G-type), and sinapyl alcohol (4) (S-type) units \(^9,10\); lignin of gymnosperms consists almost entirely of G-type lignin (G-lignin); dicotyledonous angiosperms produce a mixture of G- and S-type lignins (GS-lignin). All three types of lignin can be found in variable quantities in monocotyledonous lignin (GSH4-lignin). Incomplete or modified monolignols accompany these three main lignin types in woody materials \(^11\). The diversity in binding types found in lignin stems from its biosynthesis, in which initially formed monolignol radicals re-combine. The resulting dehydrodimer is used by the same enzymes to induce a second round of radical polymerisation, etc. hence lignin formation is not an organised living radical polymerisation, but rather a series of polymerisation termination reactions involving mainly oligomers. The relative abundance of the H/G/S units, in connection with the distribution of different motifs of interunit linkages, which result from the various coupling events, is generally used for chemically and structurally describing a lignin sample. The three monomer units form – as far as structurally possible – up to eight characteristic interunit linkages (Figure 1C). Coupling is generally favoured at the β-position of the monolignol species, resulting in the formation of arylglycerol-β-aryl ethers (β-O-4’ motif, 1f), phenylcoumarans (β-S’ motif, 1g), pinoresinols (β-b’ motif, 1h), diphenylethane dimers (β-1’ motif, 1i), and spirodienones (SD motif, 1j). Dilignols and higher oligomers preferentially couple at positions 4 and 5, yielding diaryl ethers (4-O-5’ motif, 1k), dibenzodioxocines (5,5’-αβ-O-4’ motif, 1l) and biphenyls (5-5’ motif, 1m).

Not always mentioned explicitly, but nonetheless important to keep in mind, is the fact that both the mechanical and the chemical means used for isolating the lignin from the biomass source, induce changes to the natural structure of a plant-specific lignin. Besides the industrially important lignins, Kraft-lignin, sulphite-lignin (lignosulphonate), pyrolysis lignin and steam-explosion lignin, also the milled wood lignin (MWL), acidolysis lignin (AL), cellulolytic enzyme lignin (CEL), enzymatic mild acidolysis lignin (EMAL) and

---

*University of Rome ‘Tor Vergata’, Department of Chemical Sciences and Technologies, Via della Ricerca Scientifica, 00133 Rome, Italy

* Corresponding author.
E-mail address: crestinii@stc.uniroma2.it (Claudia Crestini)

---

*Supporting information for this article is available at http://www.csbj.org*
Figure 1. (A) Representative structures of lignin biopolymers, (B) specific lignin types, and (C) main linkage motifs found in lignin and lignin extracts. See main text for details.
organismsol lignin show characteristic structural features that are reminiscent of the name-giving isolation procedure; MWL commonly serves as the basis for comparison [6,12]; in fact the oligomer-type lignin structure shown in Figure 1A was derived on the basis of a sample of milled wood lignin [13].

The lack of regular patterns other than the recurring binding motifs mentioned above, is – to the best of our current knowledge and understanding – indeed only the result of the interesting biosynthetic pathway that starts with the generation of the monolignol radicals [14,15], and proceeds with the recombination of these radicals rather than with a propagation of a classical radical chain reaction. Involving the so-generated dimers, trimers, etc. in the same reaction cycles lead then to the observed oligomeric lignin structures shown in Figure 1A.

As mentioned before, the different structural characteristics of a given lignin sample are best to be elucidated using heteronuclear 2D-NMR techniques, both with respect to qualitative and quantitative analyses. As described in greater detail by others before [4-6], the characterisation of lignin by 2D-NMR techniques has benefitted from both the advancements in NMR technologies as such, and from the advances in delineating and ascribing the NMR signals based on model studies and comparative lignin studies [4,5,6, and cited references]. Several difficulties, however, have to be considered and overcome for achieving a reliable result, especially within quantitative analyses: i) the sample has to be fully dissolvable in the solvent of choice, and stay dissolved during the time of the measurements; ii) standards have to be inert and stable over the time needed for the measurements; iii) signal to noise ratios have to be reasonable within reasonable time-frames needed for the measurement. While the solubility issue can often be resolved by simply acetylating the sample, and/or by working in DMSO-d6 as solvent, and while optimum external and internal standards have been identified long ago [6], achieving a reasonable S/N-ratio within a reasonable experiment time is still a challenge within the given analyses parameters set by sample stability, solubility etc.

Inspired by similar problems, modified heteronuclear single quantum coherence (2D-HSQC) protocols have been developed by other groups [6,16,17,18]. We have recently started to adopt these time- and S/N-optimised methodologies for the examination of MWL lignins, and were able to obtain important structural information using these optimised techniques in connection with other standard techniques used in lignin characterisation [13,19]. In fact the elucidation of the linearity of the oligomers of MWL lignin was achieved using QQ-HSQC [13,17]. Although this technique proved very useful, we set out to test another quantitative HSQC-technique, HSQC, which is simpler to implement than the QQ-HSQC method. We report here the first results obtained by the HSQC method applied in the analyses of an acetylated standard MWL sample that is well studied otherwise. We were especially interested to see whether the different pulse sequences result in generally different quantifications, or whether certain motifs were distinctly affected due to the effects of certain characteristics of the pulse sequences and acquisition methods.

**Experimental Details**

**Sample preparation**

Milled wood lignin was isolated from Norway spruce wood as reported before [13] by slight modification of the Björkman method [20]. Acetylation was performed in pyridine/acetic anhydride (V/V = 1:1) at 50 °C for 48 h. Ethanol was added, and the volatiles were removed in vacuo; the procedure was repeated twice. Then, toluene was added, and the volatiles were quickly removed in vacuo; this was repeated twice, as well. Finally, chloroform was added, and the volatiles were removed in vacuo; again, this was repeated twice. The solid was then dried overnight in vacuo (see Note-1).

![Figure 1](https://via.placeholder.com/150)

**Figure 1.** Experimental details.

**Figure 2.** QQ-HSQC contour diagrams obtained for Norway spruce MWL using a 600 MHz NMR spectrometer equipped with a cryoprobe. Signals corresponding to important lignin interunit bonding motifs are indicated (N.b: the circles do only indicate the signal positions used for integration, and do not mirror the areas used for quantitative analyses). Details concerning sample preparation, data acquisition, and data processing are given in the Experimental Details section and in reference [12].

**Measurements**

All spectra were acquired at 303 K with a Bruker Avance 600 spectrometer equipped with a cryoprobe. The sample consisted of 80 mg of acetylated lignin dissolved in 600 μL of DMSO-d6. A matrix consisting of 256 x 2048 points was obtained in eight scans. QQ-HSQC measurements were performed in accordance with the original reference [17] as reported before [13,19]. In the HSQC-related measurements, the second and third HSQC were obtained as repetitions of the basic HSQC scheme, according to the published procedure [18].
Table 1. Representative absolute intensities and deduced relative abundances (in % C9) of the internal standard (G C2-H (G2)) and important interunit bonding motifs in Norway spruce MWL. Samples were acetylated before spectral analysis (see Experimental Details). Abbreviations used in the table: abs. – absolute; rel. – relative; std. – standard; ox. – oxidised.

<table>
<thead>
<tr>
<th>entry</th>
<th>motif</th>
<th>absolute intensity</th>
<th>average abs. intensities per group/motif</th>
<th>rel. abundances [% C9]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>QQ-HSQC</td>
<td>HSQCδ</td>
<td>QQ-HSQC</td>
</tr>
<tr>
<td>1</td>
<td>G2</td>
<td>(G C2-H)</td>
<td>613</td>
<td>11731</td>
</tr>
<tr>
<td>2</td>
<td>β-O-4' (Hα of std. motif)</td>
<td>246</td>
<td>4481</td>
<td>246</td>
</tr>
<tr>
<td>3</td>
<td>β-O-4' (Hδ incl. ox. der.)</td>
<td>262</td>
<td>4712</td>
<td>262</td>
</tr>
<tr>
<td>4</td>
<td>β-O-4' (Cα-H of ox. der.)</td>
<td>38</td>
<td>617</td>
<td>38</td>
</tr>
<tr>
<td>5</td>
<td>β-5' (Hα)</td>
<td>73</td>
<td>1065</td>
<td>73</td>
</tr>
<tr>
<td>6</td>
<td>β-1' (Hδ)</td>
<td>10</td>
<td>162</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>5,5′,αβ-O-4' (Hα)</td>
<td>16</td>
<td>271</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>cinnamyl alcohol (Hγ)</td>
<td>27</td>
<td>513</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>β-β' (Hδ)</td>
<td>41</td>
<td>834</td>
<td>41</td>
</tr>
</tbody>
</table>

* Data shown are for one representative set of measurements whose contour diagrams are shown in the pictures. QQ-HSQC and HSQCδ measurements were obtained from two distinct acetylation procedures using the very same Norway spruce MWL sample as starting material.

* Values in brackets represent the relative abundance determined by subtracting the relative abundance of the Hα of the β-O-4' standard motif (1f) from that of the Hδ of the β-O-4' standard motif that also comprises the Hα abundances of the in benzylic position oxidised form of β-O-4' (motif 1o).

* Abundances have been determined by subtracting the intensity of 5,5′,αβ-O-4' (Hα) from the intensity of the combined signal of [5,5′,αβ-O-4' (Hα) + β-β' (Hδ)] (HSQCδ: δH = 4.79, δC = 84.76 ppm; QQ-HSQC: δH = 4.74, δC = 84.50 ppm).

Data processing

NMR data were processed with MestreNova (Version 8.1.1, Mestrelab Research) by using a 60°-shifted square sine-bell apodisation window; after Fourier transformation and phase correction a baseline correction was applied in both dimensions. The final matrix consisted of 1024 x 1024 points, and cross-peaks were integrated with the same software that allows the typical shape of peaks present in the spectrum to be taken into account. Extrapolations based on the values of the volumes of the peaks of interest of the three consecutive HSQC measurements were performed using MS Excel 2010, to yield the volumes of the different peaks for quantitative analyses.

Raw spectral data and table calculations used for processing are available as Supporting Information or form the authors upon request.

Note-1: We perform the acetylation as outlined in the Experimental Details as smoothly as possible, using rotary vane pumps for achieving a good vacuum that ensures a minimum of time for, and low temperatures during the evaporation of solvents within the work-up sequence, in order to avoid undesired structural changes within the lignin sample.

Results and Discussion

Quantitative CPMG-adjusted HSQC (Q-CAHSQC) measurements [16] were often too long with respect to the integrity of the samples. Quick quantitative HSQC (QQ-HSQC) measurements – as developed by Peterson and Loenig [17] – were thus adopted for measurements of samples of lignin, which we prefer to acetylate before measurement in order to ensure complete solubility of the samples in the NMR solvents (here DMSO-d6) at concentrations that should guarantee high signal-to-noise (SN) ratios (see Note-1). After careful calibration of pulses, we obtained the best signal-to-noise ratios using a 600 MHz NMR spectrometer equipped with a cryoprobe. We based quantification for softwood lignins on the guaiacyl (G) C2-H signals (G2) as internal standard (IS), since the C2 position of the guaiacyl unit never is substituted, and since it is clearly assignable in the HSQC spectrum. Routinely, though the amount of G2 directly reflects the amount of C2 units, we validated the quantification using the C2-sigal of the C2-H signals of both the para-hydroxyphenyl (H2) and the syringyl (S2) units for hardwood and grass lignins. Figure 2 shows exemplary spectra obtained for acetylated Norway spruce milled wood lignin (MWL).
Figure 3. (A) Full range HSQC, obtained for acetylated Norway spruce MWL using a 600 MHz spectrometer equipped with a cryoprobe. (B) Zoom-ins for the aromatic region ($\delta_{1H} 8.0-6.0, \delta_{13C} 135-100$), and the aliphatic region ($\delta_{1H} 6.5-2.5, \delta_{13C} 95-45$) of contour-diagrams obtained for interesting regions for the HSQC$_1$, HSQC$_2$, and HSQC$_3$. (C) Backwards extrapolation for representative signals. (D) Graphical comparison of the intensities of signals corresponding to the main interunit bonding motifs found in the Norway spruce MWL sample (as % C9); legend: A - $\beta$-$O-4'$ (H$_{\alpha}$), B - $\beta$-$O-4'$ (H$_{\beta}$ including oxidised derivative), C - $\beta$-$O-4'$ (C$_{2}$-H of oxidised derivative), D - $\beta$-$S'$ (H$_{3}$), E - $\beta$-$1'$ (H$_{3}$), F - 5,5'-$\alpha$-$\beta$-$O-4'$ (H$_{3}$), G - $\beta$-$B'$ (H$_{3}$), H - cinnamyl alcohol (H$_{3}$).
Based on our own experience, as well as on standard literature on lignin NMR [4-6,21], we decided on the integration ranges roughly indicated in Figure 2 (red cycles), and given in the Supporting Information, which led to the absolute intensity values listed in Table 1, where we also list the integration values found by the QQ-HSQC. Very reproducible results were obtained before for this method, with errors of 0.01–2 bonds per 100 C6 units [19].

Using these results as a base for further evaluations, we adopted a different sequence for quantitative H13C-HSQC studies that was developed originally by Hu, Westler, and Markley, and that consists of multiple repetitions of a basic HSQC pulse sequence [18]. Using this approach, a so-called HSQC0 is obtained by extrapolation of data collected in the aforementioned series of repetitions of a basic HSQC. The experiment starts with a standard HSQC sequence (HSQC1, Figure 3A). HSQC1 and HSQC0 are then obtained by simple double repetition of the basic scheme using incremented repetition times. All the problems associated with quantification in a standard HSQC (J-coupling modulation, T1 relaxation, etc.) produce a linear attenuation of the signals going from HSQC1 to HSQC0. Following integration of the relevant peaks in the three experiments, an HSQC0, i.e. a problem-free HSQC, is then obtained by backward extrapolation of the data. Figure 3B shows contour diagrams obtained for three consecutive HSQC spectra obtained for the same acetylated Norway spruce MWL lignin analysed before with the QQ-HSQC pulse sequence (Figure 2). Exemplary extrapolation curves obtained by integration of the signals the internal standard (G C-H (G)) and important interunit bonding motifs in Norway spruce MWL for HSQC1, HSQC0, and HSQC3 are shown in Figure 3B: Figure 3C shows the graphical analyses towards the determination of intensities in the theoretical HSQC0. In general, the quality of the linear regressions used for the backward extrapolation is acceptable (R² ≥ 0.9); only two extrapolations of signals of lower intensities, namely for the oxidised form of the interunit bonding motif β-O-4′ (1o), and for the cinnamyl endgroups (groups 2-4 (Figure 1) linked via an ether bond to a lignin backbone) result in poorer coefficients of determination (R² = 0.88, 0.87, respectively). Whereas the difficulties in the determination of the cinnamyl endgroups might result from the fact that quantification is based on a CH3=CH-group rather than a simple CH1-group, problems in the quantification of the oxidised β-O-4′ unit are likely to result from overlaps. The quality of the latter can be estimated by comparing the relative abundance with the difference between the relative abundances found for the joint signal of the H₉ of both the β-O-4′ motif (1f) and its oxidised analogue 1o, and the signal for the H₁₀ of the β-O-4′ motif (1f): for both methods, this deduction of the relative abundance of the oxidised form 1o of the β-O-4′ motif (1f) results in only half the amount. Based on the quality of the spectral data, however, one of the two ways for determining the amount of oxidised β-O-4′ units is more reliable; in the present case, the high purity of the MWL sample used renders the subtraction-method more viable.

The HSQC0-intensities are listed in Table 1 for comparison with the previously obtained data. Figure 3D represents an additional graphical comparison of the results obtained for the most important interunit bonding motifs in lignin.

The direct comparison of the data obtained using the two different methods for obtaining quantitative HSQC data reveals the same relations between important interunit bonding motifs. Furthermore, also the results of the quantifications against the internal standard (G C-H (G)) are showing essentially the same results, considering the experimental errors in the measurements for all of the motifs investigated except the β-5′ (H₉) motif. This motif, however, is known to be quantified only with difficulties. Given the difficult nature of the lignin as such, our results indicate that both quantitative HSQC methods deliver very similar, and thus comparable results, and that both can be used in lignin analyses. Comparing the two methods, with respect to the time that is actually needed for obtaining the spectra, and with respect to the plain key figures of the acquisitions, etc., it has to be clear that the easier to be implemented HSQC0 is a series of three consecutive HSQC measurements, in which each experiment has the same relaxation delay, the same number of scans and the same number of experiments as the QQ-HSQC: in our case, 16 scans were accumulated using a relaxation delay of 12 seconds over a matrix consisting of 2048 x 256 points. Thus, the total acquisition time of HSQC0 experiment is three times the experimental time of the QQ-HSQC.

Having obtained good results using the HSQC0 method on a 600MHz spectrometer equipped with a cryoprobe, we tested whether comparable results can be obtained running the HSQC0 program on a more common 400 MHz spectrometer without cryoprobe. The results we obtained using this machine were unfortunately not matching the above reported results, the main difficulties arising from very unsatisfying S/N ratios obtained within the HSQC2 and HSQC0 measurements (data not reported). We then increased the number of scans in connection using a non-uniform sampling scheme [22]. Unfortunately, also this method, at the end, did not work using a 400 MHz NMR spectrometer equipped with a standard broadband (BB0) probe. Further studies are currently under way using a 400 MHz NMR spectrometer equipped with a cryoprobe, in order to elucidate the importance of the field strength for these types of measurements.

Concluding, we found that both, QQ-HSQC and HSQC0 as standard methods for obtaining quantitative HSQC data, can be successfully exploited for lignin analyses, delivering generally comparable data, with QQ-HSQC being the faster method, while the HSQC0 method is easier to implement. Further trials aiming at reducing the time needed for obtaining a quantitative HSQC spectrum for a standard lignin sample, by using non-uniform sampling techniques in connection with the HSQC0 method, did not result in viable results.

Acknowledgements

The EU NMR facility of Utrecht is gratefully acknowledged for technical assistance in the frame of the European Union project “EU-NMR-European Network of Research Infrastructures for Providing Access & Technological Advancement in Bio-NMR” FP-2005-RII3 Contract 026145.

Citation


References


Keywords:
natural polymer, structure elucidation, quantitative heterocorrelated NMR, QQ-HSQC, HSQC0.

Competing Interests:
The authors have declared that no competing interests exist.

© 2013 Sette et al. Licensee: Computational and Structural Biotechnology Journal. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly cited.

What is the advantage to you of publishing in Computational and Structural Biotechnology Journal (CSBJ)?

- Easy 5 step online submission system & online manuscript tracking
- Fastest turnaround time with thorough peer review
- Inclusion in scholarly databases
- Low Article Processing Charges
- Author Copyright
- Open access, available to anyone in the world to download for free

WWW.CSBJ.ORG